Dr D. Sinden

CTMS-MAT-13: Numerical Methods

Problem Sheet 2 Solutions. Released: 26 February 2025

Exercise 1: Find the solution to the system of equations using Gaussian elimination:

$$\begin{aligned} -5x + 5y + 10z &= 45 \\ 3x - y - z &= -3 \\ 3x - 6y + 6z &= 15 \end{aligned}$$

Firstly, set up the augmented matrix

$$\left[\begin{array}{ccc|c} -5 & 5 & 10 & 45 \\ 3 & -1 & -1 & -3 \\ 3 & -6 & 6 & 15 \end{array}\right]$$

Step 2: Simplify row 1 by multiplying R_1 by $-\frac{1}{5}$:

$$\left[\begin{array}{ccc|c} 1 & -1 & -2 & -9 \\ 3 & -1 & -1 & -3 \\ 3 & -6 & 6 & 15 \end{array}\right]$$

Step 3: Eliminate first column $R_2 \leftarrow R_2 - 3R_1$ and $R_3 \leftarrow R_3 - 3R_1$:

$$\left[\begin{array}{ccc|c} 1 & -1 & -2 & -9 \\ 0 & 2 & 5 & 24 \\ 0 & -3 & 12 & 42 \end{array}\right]$$

Step 4: Eliminate second column $R_3 \leftarrow R_3 + \frac{3}{2}R_2$:

$$\left[\begin{array}{ccc|c} 1 & -1 & -2 & -9 \\ 0 & 2 & 5 & 24 \\ 0 & 0 & \frac{39}{2} & 78 \end{array}\right]$$

Back substitution. From row 3 deduce:

$$\frac{39}{2}z = 78 \rightarrow z = 4$$

From row 2:

$$2y + 5 \times 4 = 24 \rightarrow 2y = 4 \rightarrow y = 2.$$

From row 1:

$$x - 2 - 2 \times 4 = -9 \rightarrow x = 1.$$

Thus

$$x = 1, \quad y = 2, \quad \text{and} \quad z = 4.$$

Exercise 2: Let

$$A = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 3 & 2 & 1 & 4 \\ 5 & 8 & 6 & 3 \\ 4 & 2 & 5 & 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}.$$

a) Check if Gaussian elimination can be applied to solve $A\boldsymbol{x} = \boldsymbol{b}$.

b) Solve $A\boldsymbol{x} = \boldsymbol{b}$ for $\boldsymbol{x}$ by Gaussian elimination with scaled partial pivoting.

a) Check if Gaussian elimination can be applied Gaussian elimination can be applied if the matrix is non-singular (i.e., all pivots are non-zero, possibly after row exchanges). Form the augmented matrix and perform elimination:

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 1 \\ 3 & 2 & 1 & 4 & 1 \\ 5 & 8 & 6 & 3 & 1 \\ 4 & 2 & 5 & 3 & -1 \end{array}\right]$$

Step 1: Eliminate below pivot $a_{11} = 1$

$$R_2 \leftarrow R_2 - 3R_1,$$
$$R_3 \leftarrow R_3 - 5R_1,$$
$$R_4 \leftarrow R_4 - 4R_1.$$

Thus

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 1 \\ 0 & 5 & -5 & 1 & -2 \\ 0 & 13 & -4 & -2 & -4 \\ 0 & 6 & -3 & -1 & -5 \end{array}\right].$$

Step 2: Pivot $a_{22} = 5 \neq 0$. Eliminate below:

$$R_3 \leftarrow R_3 \frac{13}{5} R_2,$$
$$R_4 \leftarrow R_4 - \frac{6}{5} R_2.$$

Augmented matrix is now

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 1 \\ 0 & 5 & -5 & 1 & -2 \\ 0 & 0 & 9 & -\frac{23}{5} & \frac{6}{5} \\ 0 & 0 & 3 & -\frac{11}{5} & -\frac{13}{5} \end{array}\right]$$

Step 3: Pivot $a_{33} = 9 \neq 0$. Eliminate below: $R_4 \leftarrow R_4 - \frac{1}{3} R_3$ so finally

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 1 \\ 0 & 5 & -5 & 1 & -2 \\ 0 & 0 & 9 & -\frac{23}{5} & \frac{6}{5} \\ 0 & 0 & 0 & -\frac{2}{3} & -3 \end{array}\right]$$

All pivots are non-zero: $1, 5, 9, -\frac{2}{3}$. Therefore, Gaussian elimination can be applied.

b) With scaled partial pivoting, at each step choose the row with the largest ratio $\frac{|a_{ik}|}{s_i}$ as pivot, where s_i is the scale factor (max absolute value in row i of the *original* matrix).

Scale factors:

$$s_1 = \max(|1|, |-1|, |2|, |1|) = 2,$$
$$s_2 = \max(|3|, |2|, |1|, |4|) = 4,$$
$$s_3 = \max(|5|, |8|, |6|, |3|) = 8,$$
$$s_4 = \max(|4|, |2|, |5|, |3|) = 5.$$

so that

$$\vec{s} = (2, 4, 8, 5).$$

Choose the pivot for column 1. Ratios are $\frac{|1|}{2} = 0.5$, $\frac{|3|}{4} = 0.75$, $\frac{|5|}{8} = 0.625$, $\frac{|4|}{5} = 0.8$. As the maximum is 0.8, so swap rows 1 and 4, to yield

$$\left[\begin{array}{cccc|c} 4 & 2 & 5 & 3 & -1 \\ 3 & 2 & 1 & 4 & 1 \\ 5 & 8 & 6 & 3 & 1 \\ 1 & -1 & 2 & 1 & 1 \end{array}\right].$$

Eliminate below pivot, to give

$$\left[\begin{array}{cccc|c} 4 & 2 & 5 & 3 & -1 \\ 0 & \frac{1}{2} & -\frac{11}{4} & \frac{7}{4} & \frac{7}{4} \\ 0 & \frac{11}{2} & -\frac{1}{4} & -\frac{3}{4} & \frac{9}{4} \\ 0 & -\frac{3}{2} & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} \end{array}\right].$$

Note that now $s = (5, 4, 8, 2)$

Now eliminate entries from the second column. To see whether rows need to be swapped, consider the ratios = $(|\frac{1}{2}|/4, |\frac{11}{2}|/8, |\frac{3}{2}|/2) = (1/8, 11/16, 3/4)$. As the largest is row 4, so swap $R_2 \leftrightarrow R_4$:

$$\left[\begin{array}{cccc|c} 4 & 2 & 5 & 3 & -1 \\ 0 & -\frac{3}{2} & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} \\ 0 & \frac{11}{2} & -\frac{1}{4} & -\frac{3}{4} & \frac{9}{4} \\ 0 & \frac{1}{2} & -\frac{11}{4} & \frac{7}{4} & \frac{7}{4} \end{array}\right].$$

Eliminate below pivot:

$$\left[\begin{array}{cccc|c} 4 & 2 & 5 & 3 & -1 \\ 0 & -\frac{3}{2} & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} \\ 0 & 0 & \frac{5}{2} & \frac{1}{6} & \frac{41}{6} \\ 0 & 0 & -\frac{5}{2} & \frac{11}{6} & \frac{13}{6} \end{array}\right].$$

To eliminate the last entry of the third column, consider the ratios = $(|\frac{5}{2}|/8, |\frac{5}{2}|/4) = (5/16, 5/8)$. Again, largest is row 4, so swap $R_3 \leftrightarrow R_4$:

$$\left[\begin{array}{cccc|c} 4 & 2 & 5 & 3 & -1 \\ 0 & -\frac{3}{2} & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} \\ 0 & 0 & -\frac{5}{2} & \frac{11}{6} & \frac{13}{6} \\ 0 & 0 & \frac{5}{2} & \frac{1}{6} & \frac{41}{6} \end{array}\right]$$

and eliminate below the pivot value to yield

$$\left[\begin{array}{cccc|c} 4 & 2 & 5 & 3 & -1 \\ 0 & -\frac{3}{2} & \frac{3}{4} & \frac{1}{4} & \frac{5}{4} \\ 0 & 0 & -\frac{5}{2} & \frac{11}{6} & \frac{13}{6} \\ 0 & 0 & 0 & 2 & 9 \end{array}\right].$$

Back substitution:

$$\begin{aligned} 2x_4 = 9 &\Rightarrow x_4 = \tfrac{9}{2}, \\ -\tfrac{5}{2}x_3 + \tfrac{11}{6}x_4 = \tfrac{13}{6} &\Rightarrow x_3 = \tfrac{73}{30}, \\ -\tfrac{3}{2}x_2 + \tfrac{3}{4}x_3 + \tfrac{1}{4}x_4 = \tfrac{5}{4} &\Rightarrow x_2 = \tfrac{17}{15}, \\ 4x_1 + 2x_2 + 5x_3 + 3x_4 = -1 &\Rightarrow x_1 = -\tfrac{217}{30}. \end{aligned}$$

Thus

$$\vec{x} = \left(-\frac{217}{30}, \frac{17}{15}, \frac{73}{30}, \frac{9}{2}\right).$$

Exercise 3: Let

$$A = \begin{pmatrix} 4 & 1 & 2 \\ 1 & 5 & 1 \\ 2 & 1 & 6 \end{pmatrix}.$$

a) Show that matrix A is positive definite.

b) Compute the LU decomposition of A where L has ones on the leading diagonal.

c) Compute the Cholesky decomposition of A.

a) One way of showing that a matrix A is positive definite: a symmetric matrix is positive definite if all its leading principal minors are positive (Sylvester's criterion). It is often easier than finding the eigenvalues. First, note that A is symmetric: $A = A^T$

Leading principal minors:

$$\begin{aligned} M_1 &= 4 > 0 \\ M_2 &= \begin{vmatrix} 4 & 1 \\ 1 & 5 \end{vmatrix} = 4(5) - 1(1) = 20 - 1 = 19 > 0 \\ M_3 &= \det(A) = \begin{vmatrix} 4 & 1 & 2 \\ 1 & 5 & 1 \\ 2 & 1 & 6 \end{vmatrix} \end{aligned}$$

Expanding along the first row:

$$\begin{aligned} &= 4\begin{vmatrix} 5 & 1 \\ 1 & 6 \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ 2 & 6 \end{vmatrix} + 2\begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} \\ &= 4(30 - 1) - 1(6 - 2) + 2(1 - 10) \\ &= 4(29) - 4 + 2(-9) = 116 - 4 - 18 = 94 > 0 \end{aligned}$$

Since all leading principal minors of the symmetric matrix are positive, A is positive definite.

Alternatively show all eigenvalues are real and positive. The characteristic polynomial is

$$\lambda^3 - 15\lambda^2 + 68\lambda - 94 = 0.$$

Numerically solving this gives $\lambda_1 \approx 2.7076$, $\lambda_2 \approx 4.3973$ and $\lambda_3 \approx 7.8951$.

It is possible to solve this analytically. Let $\lambda = t + 5$ to get the depressed cubic $t^3 - 7t - 4 = 0$, which has the closed form solution

$$\lambda_k = 5 + 2\sqrt{\frac{7}{3}}\cos\left(\frac{1}{3}\cos^{-1}\left(\frac{6\sqrt{3}}{7\sqrt{7}}\right) - \frac{2\pi k}{3}\right), \quad k = 0, 1, 2.$$

b) When trying to find $A = LU$, note that in general this is under-determined. Thus, to find a unique solution, it is enforced that L is lower triangular with 1's on the diagonal and U is upper triangular.

Evaluating $a_{11} = 1 \cdot u_{11} \Rightarrow u_{11} = 4$. Similarly, $u_{21} = 1$ and $u_{31} = 2$.

Now, knowing u_{11} then, the next step is to equate the values in the first column of A. This yields $l_{21} = \frac{1}{4}$ and $l_{31} = \frac{2}{4} = \frac{1}{2}$.

Equating values of the second column of A first gives $5 = 1/4 + u_{22}$, thus $u_{22} = 19/4$. Next, $1 = 2/4 + u_{32}$ yields $u_{32} = 1/2$.

Continuing for the third row of A, then gives $l_{32} = \frac{1/2}{19/4} = \frac{2}{19}$ After elimination:

$$u_{33} = 5 - \frac{2}{19} \cdot \frac{1}{2} = 5 - \frac{1}{19} = \frac{94}{19}$$

Hence:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{1}{2} & \frac{2}{19} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 4 & 1 & 2 \\ 0 & \frac{19}{4} & \frac{1}{2} \\ 0 & 0 & \frac{94}{19} \end{pmatrix}.$$

c) Since A is positive definite and symmetric, we can write $A = \tilde{L}\tilde{L}^T$ where $\tilde{L}$ is lower triangular with positive diagonal entries. Let

$$\tilde{L} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix},$$

then by equating coefficients of A to $\tilde{L}$ using $A = \tilde{L}\tilde{L}^T$ yields, for column 1:

$$\begin{aligned} l_{11}^2 &= a_{11} = 4 \Rightarrow l_{11} = 2, \\ l_{21}l_{11} &= a_{21} = 1 \Rightarrow l_{21} = \frac{1}{2}, \\ l_{31}l_{11} &= a_{13} = 2 \Rightarrow l_{31} = 1. \end{aligned}$$

Column 2:

$$\begin{aligned} l_{21}^2 + l_{22}^2 = 5 &\Rightarrow \frac{1}{4} + l_{22}^2 = 5 \Rightarrow l_{22} = \frac{\sqrt{19}}{2}, \\ l_{31}l_{21} + l_{32}l_{22} = 1 &\Rightarrow \frac{1}{2} + l_{32} \cdot \frac{\sqrt{19}}{2} = 1 \Rightarrow l_{32} = \frac{\sqrt{19}}{19} = \frac{1}{\sqrt{19}}. \end{aligned}$$

Column 3:

$$\begin{aligned} l_{31}^2 + l_{32}^2 + l_{33}^2 &= 6 \\ 1 + \frac{1}{19} + l_{33}^2 &= 6 \Rightarrow l_{33}^2 = \frac{94}{19} \Rightarrow l_{33} = \sqrt{\frac{94}{19}}. \end{aligned}$$

Thus

$$\tilde{L} = \begin{pmatrix} 2 & 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{19}}{2} & 0 \\ 1 & \frac{1}{\sqrt{19}} & \sqrt{\frac{94}{19}} \end{pmatrix}.$$

Exercise 4:

a) Using

$$\sum_{k=1}^{n-1} k\,(k+1) = \sum_{k=1}^{n-1} k^2 + \sum_{k=1}^{n-1} k$$
$$= \frac{1}{6}(n-1)n(2n-1) + \frac{1}{2}n(n-1),$$

or otherwise, explain why, for a $(n \times n)$-matrix, Gaussian elimination has $\mathcal{O}\left(n^3\right)$ complexity.

b) Explain why, for a $(n \times n)$-tridiagonal matrix, Gaussian elimination only has $\mathcal{O}\,(n)$ complexity. (This is known as the Thomas algorithm)

a) At the k-th stage of the first part of Gaussian elimination, it necessary to perform one division, $(n-k+1)$ multiplications, and $(n-k+1)$ subtractions to eliminate an entry from the j-th row.

But at the k-th stage there are a total of $n-k$ rows to remove entries to create the triangular form. This means in total $n-k$ divisions and $(n-k)(n-k+1)$ multiplications and subtractions.

Rewrite the summation as

$$\sum_{k=1}^{n-1}(n-k) = \sum_{k=1}^{n-1} n - \sum_{k=1}^{n-1} k = n(n-1) - (n-1)n/2 = \frac{n(n-1)}{2}.$$

Therefore,

$$\sum_{k=1}^{n-1}(n-k) = \frac{1}{2}n\,(n-1)$$

divisions and, applying the same procedure show

$$\sum_{k=1}^{n-1}(n-k)\,(n-k+1) = \frac{1}{3}n\left(n^2-1\right)$$

multiplications and the same number of subtractions.

For back substitution, to obtain each value, one division, $n-k$ multiplications and $n-k$ subtractions. Thus, for back-substitution, there are, in total, n divisions and

$$\sum_{k=1}^{n}(n-k) = \frac{1}{2}n(n-1)$$

multiplications and the same number of subtractions which must be performed.

Thus, for Gaussian elimination, this gives $n(n+1)/2$ divisions, and both $(2n^3+3n^2-5n)/6$ for both multiplication and division.

b) Consider $Ax = d$, and let the tridiagonal matrix be

$$A = \begin{pmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & a_3 & b_3 & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ 0 & & & a_n & b_n \end{pmatrix}$$

The first equation of the linear system is $b_1x_1 + c_1x_2 = d_1$ has two unknowns. The second equation has three unknowns $a_2x_1 + b_2x_2 + c_2x_3 = d_2$. Write

$$a_ix_{i-1} + b_ix_i + c_ix_{i+1} = d_i \quad \text{for } i = 2, \ldots, n-1$$

and the last equation only has two unknowns $a_nx_{n-1} + b_nx_n = d_n$.

For the second equation, the unknown x_1 can be eliminated via the row operation $r_2 \mapsto b_1 r_2 - a_2 r_1$, yielding an equation in two unknowns

$$(b_1 b_2 - c_1 a_2)\, x_2 + c_2 b_1 x_3 = d_2 b_1 - d_1 a_2.$$

This procedure can be repeated, eliminating unknowns from each row (by eliminating every a_i), so that the last row will only have a single unknown. This gives a upper triangular matrix.

Then solving for the x_n means that will be possible to solve for x_{n-1}.

Both the forward elimination and back substitution require $\mathcal{O}\left(n\right)$ operations. The forward elimination has one division, one multiplication and one subtraction per row, i.e. $\mathcal{O}(1)$ per row, which is $\mathcal{O}\left(n\right)$. The same operations are applied for back substitution.